Published in: Proceedings of Corpora and Tools in Linguistics, Languages, and Speech. Strasbourg, France. 3-5 Jul 2013. - Strasbourg: Université de Strasbourg, 2013. Pp. 32-34.

DESIGNING A BILINGUAL SPEECH CORPUS FOR FRENCH AND GERMAN LANGUAGE LEARNERS

Jürgen Trouvain[1], Yves Laprie[2], Bernd Möbius[1], Bistra Andreeva[1], Anne Bonneau[2], Vincent Colotte[2], Camille Fauth[2], Dominique Fohr[2], Denis Jouvet[2], Odile Mella[2], Jeanin Jügler[1], Frank Zimmerer[1]

[1]*Saarland University*

[2]*LORIA, Nancy*

Keywords: speech corpus, French-German, phonetics, language learning

1. Introduction

It can be taken for granted that the first language (L1) influences the target language to be learned (L2) on all linguistic levels including the lexicon, morphosyntax, pragmatics and, most pertinent to the proposed corpus, sound structure and its phonetic implementation (e.g. Flege & Davidian 1984, Flege 1995). A typical example of L1-L2 interference on the segmental level is the realization of stop consonants. The phonological [voice] contrast for stops in French is phonetically realized by voicing, whereas in German this is mainly done by aspiration, i.e. longer voice onset time. In addition, obstruents in syllable- final position are subject to devoicing in German, but not in French.

Phonetic L2 research is generally biased toward studying segmental phenomena, and most studies consider English as the target language. This is reflected in current L2 speech models, e.g. Flege's Speech Learning Model (Flege 1995) or Best's Perceptual Assimilation Model (Best 1995), which neglect prosodic phenomena. However, prosody is arguably the most salient feature for native speakers listening to nonnative speech (Jilka 2007). The location and phonetic realization of lexical stress is one of the main difficulties for learners of German as a foreign language (Hirschfeld 1994). Particularly for learners with French as L1 the deviation from a fixed position can result in considerable problems for speech production and perception, a phenomenon sometimes described as 'stress deafness' (Dupoux et al. 2001, but see counter-evidence in Schwab & Llisterri 2011).

We present the design of a corpus of native and non-native speech for the language pair French-German, with a special emphasis on the phonetic and prosodic aspects. To our knowledge there is no suitable corpus, in terms of size and coverage, currently available for the target language pair. Our design of such a corpus has three aims. (i) We will bring to the research community two non-native corpora for the French- German language pair, whereas most studies have focused on English (cf. website on "Learner corpora around the world"). (ii) The corpus will be informed by in-dept phonetic knowledge to predict the types of errors made by French and German learners. (iii) The corpus can further be used by the research community for the recognition of non-native speech, which is notoriously difficult (see e.g. Goronzy, Sahakyan & Wokurek 2001, van Doremalen, Strik & Cucchiarini 2009, Bouselmi, Fohr & Illina 2011)

2. Design

To select the target L1-L2 interference phenomena we plan to prepare a small preliminary corpus, which will be analyzed for coverage and cross-checked jointly by French and German experts. Based on this analysis, target phenomena on the phonetic and phonological level will be selected on the basis of the expected degree of deviation from the native performance and the frequency of occurrence.

In general, target phenomena will be tested in contexts of words, short phrases and sentences. Therefore, the material will cover, for both languages: (i) a phonetically rich design comprising all phonemes in relevant contexts, to achieve a reliable assessment of the entire phonemic inventory for each speaker; (ii) the most important phenomena in the phonetics and prosody of French and German as a foreign language, respectively (e.g., vowel quantity, consonantal articulation and word stress); (iii) phonological processes and alternations (e.g., final devoicing); (iv) minimal pairs.

We also plan to record utterances in a speaking style that is less prepared than reading aloud, to represent a larger range of variability. A relatively small set of recordings will be sufficient to make comparisons across levels of proficiency, and elicitation will be restricted to simple techniques such as repetition and picture naming tasks. We estimate the required total number of sentences for each language to be about 100 sentences, with a mean duration of about 3 seconds, to achieve a sufficient coverage.

3. Subjects and recordings

We plan to record a total number of 100 subjects in their foreign language (see Table 1). We include beginners as well as advanced second language learners. Among the beginners we aim at teenage learners with two to three years of L2 instruction in school, and university students. Speaker sex will be balanced in each of the six groups.

Table 1: Groups of subjects pooled across L1 (F=French, G=German), level of proficiency and age range.

# subjects	L1	L2	level	age	# subjects	L1	L2	level	age
20	F	G	beginners	18-30 yrs	20	G	F	beginners	18-30 yrs
20	F	G	advanced	18-30 yrs	20	G	F	advanced	18-30 yrs
10	F	G	beginners	15-16 yrs	10	G	F	beginners	15-16 yrs

High-quality recordings will be made in both countries, France and Germany, with highly similar acoustic characteristics (quiet room, microphones, recording devices, storage). Recording sessions will be carefully monitored for consistent quality.

4. Annotation

Annotation will be performed in two steps. (i) The entire corpus will be automatically segmented and annotated by speech-text alignment tools (e.g. Rapp 1995, Jouvet et al.2011, Fohr & Mella 2012). (ii) The entire corpus will be checked with respect to the orthographic transcription. Then a part of the corpus (approx. 10%) will be manually

checked at the levels of phones and syllables (phonetic transcription). Corrections will be made if necessary. The part of the corpus with high-quality manual annotation will be prepared as a searchable database and made available for the scientific community after completion of the project.

5. Conclusions

A bilingual speech corpus consisting of speech of French and German language learners represents an optimal starting point for detailed phonetic and phonological analyses on the segmental as well as on the prosodic level. In addition, such a phonetic learner corpus provides a rich source for first-hand information and illustration for the interested public, e.g. foreign language teachers. This corpus is also the prerequisite for the development of computer-assisted language learning software by adapting content, feedback and exercises to individual learners in the speech dimension of a foreign language.

References

Best, C.T. (1995). A direct realist view of cross-language speech perception. In: W. Strange (ed) Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research, pp. 171-206. Timonium: York Press.

Bouselmi, G., Fohr, D. & Illina, I. (2011). Multilingual recognition of non-native speech using acoustic model transformation and pronunciation modeling. International Journal of Speech Technology 15(2), pp. 203-213.

Dupoux, E., Peperkamp, S. & Sebastián-Galles, N. (2001). A robust method to study stress "deafness". Journal of the Acoustical Society of America 110, pp. 1606-1618.

Flege, J.E. & Davidian, R. (1984). Transfer and developmental processes in adult foreign language production. Journal of Applied Psycholinguistic Research 5, pp. 323-347.

Flege, J.E. (1995). Second language speech learning: Theory, findings and problems. In: W. Strange (ed) Speech Perception and Linguistic Experience: Theoretical and Method- ological Issues in Cross-Language Speech Research, pp. 233-272. Timonium: York Press.

Fohr, D. & Mella, O. (2012). CoALT: A software for comparing automatic labelling tools. Proc. Conference on Language Resources and Evaluation (LREC), Istanbul.

Goronzy, S., Sahakyan, M. & Wokurek, W. (2001). Is non-native pronunciation modeling necessary? Proc. Eurospeech, Aalborg, pp. 309-312.

Hirschfeld, U. (1994). Untersuchungen zur phonetischen Verständlichkeit Deutschlernender. (Forum Phoneticum 57). Frankfurt/M.: Hector.

Jilka, M. (2007). Different manifestations and perceptions of foreign accent in intonation. In: J. Trouvain & U. Gut (eds.) Non-Native Prosody - Phonetic Description and Teaching Practice. Berlin: Mouton De Gruyter, pp. 77-96.

Jouvet, D., Mesbahi, L., Bonneau, A., Fohr, D., Illina, I. & Laprie, Y. (2011). Impact of pronunciation variant frequency on automatic non-native speech segmentation. Proc. 5th Language & Technology Conference (LTC'11), Poznan, pp. 145-148.

Rapp, S. (1995). Automatic phonemic transcription and linguistic annotation from known text with Hidden Markov models - An aligner for German. Proc. ELSNET Goes East and IMACS Workshop "Integration of Language and Speech in Academia and Industry", Moscow.

Schwab, S. & Llisterri, J. (2011). Are French speakers able to learn to perceive lexical stress contrasts? Proc. Inter'l Congress of Phonetic Sciences (ICPhS), Hong Kong, pp. 1774-1777.

van Doremalen, J., Strik, H. & Cucchiarini, C. (2009). Optimizing non-native speech recognition for CALL applications. Proc. Interspeech, Brighton, pp. 592-595.

Website "Learner corpora around the world", maintained by Diane Goossens and Sylviane Granger at Université Catholique de Louvain, Centre for English Corpus Linguistics: https://www.uclouvain.be/en-cecl-lcworld.html [access: 25 Feb 2013]